AD-A079 236 ARMY PERSONNEL RESEARCH OFFICE WASHINGTON DC F/G 5/9
ACCURACY AND CERTITUDE IN THE DISCRIMINATION OF VISUAL NUMBER. (U)
JUN 64 R ANDREWS , F VICINO , S RINGEL

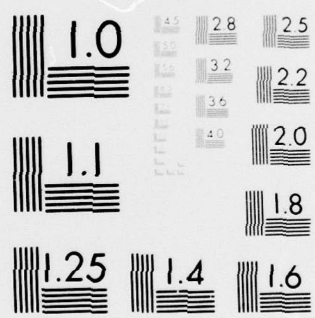
UNCLASSIFIED APRO-RM-64-6 NL

| OF |

AD A079236L

END

2-80



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

LEVEL II

Research Memorandum 64-6

ACCURACY AND CERTITUDE IN THE DISCRIMINATION OF VISUAL NUMBER

June 1964

U. S. ARMY PERSONNEL RESEARCH OFFICE

14 APRO-RM-64-6

Army Project Number
2J20901A723

Command Systems d-12

9 Research memo.,

Research Memorandum 64-6

6 ACCURACY AND CERTITUDE IN THE DISCRIMINATION OF VISUAL NUMBER .

10 Robert Andrews, Frank Vicino, ~~and~~ Seymour Ringel

12 22

Submitted by:
Joseph Zeidner
Chief, Support Systems
Research Laboratory

Approved by:
J. E. Uhlaner
Director, Research
Laboratories

11 June [redacted]64

Research Memorandums are informal reports on technical research problems. Limited distribution is made, primarily to personnel engaged in research for the U. S. Army Personnel Research Office.

040 650

ABSTRACT

ACCURACY AND CERTITUDE IN THE DISCRIMINATION OF VISUAL NUMBER

THE PROBLEM

The COMMAND SYSTEMS Task is conducting several projects to determine how accuracy of information assimilation from displays of the type used in tactical operation centers and the certitude the viewer has about this accuracy vary jointly and separately as a function of the manipulation of various information presentation variables.

There is little in the literature bearing directly on this area of inquiry, although several studies over the years have involved numerousness (discrimination of visual number) for tachistoscopically presented stimuli in which performance data, and sometimes certitude data, were obtained (Kaufman, Lord, Reese, and Volkman, 1949; Minturn and Reese, 1951; Saltzman and Garner, 1948; Taves, 1941). Fields of dots seem to be the most popular stimuli for such studies; concentric circles, numerals, and colors have been used. Generally, the functions characterizing the relationship between estimated and actual number, between errors in estimated number and actual number, and between certitude and actual number have been similar across studies. However, the majority of the studies have not included measures of certitude, and none to the author's knowledge have directly compared errors in the judgment of number with certitude about this judgment.

A common finding from these studies seems to be that estimates of number of things presented and certitude plotted as a function of the actual number of things presented are in effect discontinuous functions with the break in continuity occurring between 6 and 8 things presented. Such a finding has given rise to the postulation of at least two distinct mechanisms: subitizing (immediate apprehension) and estimating. A third category, counting, is excluded as a mechanism in such brief exposures. ABSTRACT

The extent to which reported research can provide insights into the processes involved in more complex information extraction and differing exposure times is not known. Nor is it known whether the same people would excel in performance under varying conditions or whether the certitude-performance relationship would generalize across information extraction tasks. While the determination of this information is not an integral part of the COMMAND SYSTEMS Task research program, such information could be of considerable value in understanding the nature of the psychological processes involved in extraction or assimilation of information from displays.

A brief exploratory probe was conducted as an aid in deciding whether more intensive study in this area would produce sufficiently beneficial results to justify the effort. The study was conducted prior to a nontachistoscopic experiment in which certitude and accuracy of information assimilation from visual displays as a function of amount of information presented and removed in slide updating were studied. Since the same subjects were used in the succeeding nontachistoscopic experiment, the present study related performance across tasks to the extent feasible.

PURPOSE

The general purpose of this study was a preliminary determination of the degree to which findings reported in the literature regarding discrimination of visual number and the psychological mechanisms related thereto are sustained for certain variables of interest in Command Systems research. More specifically, the objectives were:

1. to determine, for tachistoscopically presented symbols, how discrimination of visual number and subjective feelings of certitude about that discrimination are affected by (1) number of symbols displayed (4 to 22 per slide), (2) exposure time (200 and 500 milliseconds), and (3) order of exposure time (200 ms before 500 ms vs 500 ms before 200 ms).

2. to determine how the relationship of degree of error to certitude varies as a function of variations in amount (number of symbols), exposure time, and order of exposure time.

3. to compare findings of the above analysis for tachistoscopic presentations with findings for nontachistoscopic presentations using the same subject.

METHOD

EXPERIMENTAL DESIGN

Independent variables investigated in the present study were total number of symbols in the slide (amount), duration of exposure of the slide (exposure time), and order of presentation at two exposure times (order).

Nine amount levels--4, 5, 6, 7, 8, 10, 14, 18, and 22 elements per slide--were randomly presented. In order A, the nine slides were presented at 500 ms duration of exposure followed by the same nine slides in different order and orientation (upside down) at 200 ms duration. In order B, the subjects were shown the slides at 500 ms first. Subjects were randomly divided into two order groups each of which received all possible combinations of amount and exposure times.

SUBJECTS

Subjects were 32 male college graduates having normal or corrected normal vision. These subject criteria are consistent with the characteristics typical of Army officers operating in a Tactical Operations Center. Since the study did not deal with stimulus material that required military sophistication and the variables were primarily perceptual, subjects were chosen from professional and technical personnel at USAPRO. In future studies in this series as military material is included in the stimulus continuum, military subjects will be included in the sample.

STIMULUS MATERIAL

The slides used as stimulus material were 35-mm negative transparencies containing a number of identical flag symbols which appeared as white-line drawings on a dark background (⊠). The symbol is one commonly used to identify infantry units on military maps.

A 130-cell grid was used to randomly determine the location of the flag symbols for each amount level. Obvious and meaningful patterns were avoided. The symbols when projected on the screen were approximately 2-1/3 x 1-3/8 inches. To control at least in part for "location" effects (any facility or difficulty in the task due to the location of the symbols in the slide), each slide was systematically presented at one time right side up and at the next, upside down. This procedure was accomplished without loss in symbol recognition, since the symbols are preceptually the same when subjected to reorientation. The various orientations which allowed the symbols to appear in different areas of the slide increased the economy of slide production and the generalization value of the findings.

The slide projector used had a 300-watt lamp with a 3" wide-angle lens. The slides with projected image area 4' x 6' were rear projected on a 6' x 8' screen which was bisected vertically and horizontally by crosshairs. The subject's work area was illuminated by an indirect overhead lamp, rheostatically adjusted to insure enough light to read the response sheet, to reduce the possibility of after-images, and yet not interfere with the presented image. Based on MacBeth Illuminometer readings, the level of illumination on the work surface was 1.5 foot candles. The average luminance value of the indicated symbols on the screen was .4 foot lambert. These levels were maintained across all subjects and all conditions.

EXPERIMENTAL PROCEDURES

The 32 subjects were randomly divided into eight groups of four subjects each. The four subjects in each group were seated 15 feet from the viewing screen, and slide material was presented simultaneously to all four. This procedure permitted group data collection and provided some simulation of group display practices in a Tactical Operations Center.

All subjects were given the following instructions regarding the task:

> "In this first part of today's session, you will be shown a series of slides, one at a time. Each slide will be flashed on the screen for only a moment--actually less than 1 second. You are to watch the screen as closely as possible in order to determine how many flag symbols are shown on a particular slide. While waiting for a slide to appear, you should focus your attention at the intersection of crosshairs on the screen. The center of each slide shown will be at that point. After a slide has been presented, you are to write on your answer sheet the number of

flag symbols you think were on the slide. Then make a check mark in one of the five scale rectangles on your answer sheet to show how certain or uncertain you feel about the correctness of your answer. When you are finished, turn the page to the next answer sheet and wait until the experimenter announces "next slide". Again focus your attention on the center of the screen until the next slide appears, after which you fill out the answer sheet for slide number 2 and then prepare for another slide. This procedure will be followed until you have viewed a total of 18 slides, and filled out 18 answer sheets. I will now present a sample slide to give you a better idea of your task. If you have any questions, please raise your hand."

After the instructions were read (subjects also had a printed copy of the instructions before them to refer to), a practice trial was given and subjects had an opportunity to ask questions. When the subjects appeared to have adequately understood the instructions, the test session began. The testing session consisted of the presentation of 18 slides. Half the subjects received 9 slides, one at each amount level, at 500 ms (order A) followed by the same 9 slides, presented upside down, at 200 ms. The remaining 16 subjects received the first 9 slides at 200 ms followed by the same slides upside down at 500 ms (order B). The sequence of presentation of the slide material was independently randomized for each of the groups. The five-category scale was similar to that reportedly used by previous experimenters (Kaufman et al, 1949; Taves, 1941). It was end-anchored only with "absolutely uncertain" on the left end and "absolutely certain" on the right. Total elapsed time for each test session was approximately 15 minutes.

STATISTICAL ANALYSIS

Measures of dependent variables obtained were:

1. Error Score. The relative error score took into account the absolute difference between the subject's estimate of the number of symbols in a slide and the actual number of symbols presented in the slide, according to the formula:

$$\text{Error Score} = \frac{\text{True number - estimated number}}{\text{True Number}} \times 100$$

2. Estimated number of symbols.

3. Certitude. Ordinal score (1 - 5) on the certitude scale.

The error scores were used as the basic data for the analysis of variance. The data collected from the five-point scale on how certain or uncertain the subjects felt about the correctness of the answer were similarly analyzed. In addition, the degree of relation between certitude and error score was estimated by correlational analysis.

RESULTS

The distribution of error scores was found to be positively skewed, and logarithmic transformation was performed to normalize the distribution. Results of the analysis of variance on the transformed error scores are presented in Table 1. Significant differences in performance were associated with amount of symbols presented and exposure time. Neither order effects nor any of the interaction effects were found to be significant.

Table 1

ANALYSIS OF VARIANCE FOR ERROR SCORES (TRANSFORMED)

Source of Variation	d.f.	Mean Square	F	P
Order of presentation (O)	1	1.33	.99	N.S.
Btw S's within order S (O)	30	1.35		
Exposure Time (T)	1	2.53	11.00	.01
T x O	1	.07	.30	N.S.
T x S (O)	30	.23		
Amount (Number of symbols) (A)	8	8.01	28.61	.01
A x O	8	.23	.82	N.S.
A x S (O)	240	.28		
A x T	8	.41	1.46	N.S.
A x T x O	8	.19	.68	N.S.
A x T x S (O)	240	.28		
TOTAL	575			

AMOUNT (NUMBER OF SYMBOLS) AND PERFORMANCE

The nature of the statistically significant differences associated with number of symbols per slide is depicted in Figure 1 where mean error score is plotted for each amount level for the two exposure times. Except for seemingly inappropriately high error scores for 6 and 7 symbols in the 200 ms condition, increase in symbol amount produced an increase in error score. The apparent discontinuity in the curve at the 6- and the 7-symbol amounts seems to interrupt an otherwise positive relationship between number of symbols and error score. Error score at both 6 and 7 is greater than at 8; and at 7 is almost as great as at 10. When the 6- and 7-symbol slides were examined closely, the symbols appeared more widely separated than in other slides in the series. Using average distance of the symbols from fixation point (center of slide) as a dispersion score, the 7-symbol slide was found to have an average dispersion approximately 1-1/2 times greater than in the 8-symbol slide, a larger dispersion than for any of the other slides except the 18-symbol slide. Likewise, the 6-symbol slide had greater dispersion than any of the other slides except the 7- and 18-symbol slide. This dispersion increases the probability of more of the slide being subject to peripheral vision. Fixation time or time of saccadic movement between two fixations has been found to average 195 ms (Woodworth and Schlosberg, 1956, page 502). Therefore, in the 200 ms condition there was time for only one fixation, and any stimulus peripheral to the foveal image would remain peripheral and thus more difficult to detect. The poor performance in the 7-symbol condition could have been a function of this circumstance. If so, the 500 ms condition (where there was time for two fixations) would not have suffered as much by the greater symbol dispersion. An examination of Figure 1 supports this hypothesis, and thus tends to confirm that the rather inconsistently high error score at 7 symbols may have been a function of an unfortunate choice of slide to represent the 7-symbol amount. Though less pronounced, the slightly smaller error at 22 symbols than at 18 may be attributable to the same cause. To summarize the relationship between amount and error score, as the number of symbols per slide increased, error score tended to increase, from approximately 5-10% error at 4 symbols to 25-30% error at 22 symbols.

Figure 2 presents the median number of symbols estimated by the 32 subjects (separately for each exposure time) for each symbol amount presented. The dotted line represents the actual number of symbols presented. When 4 symbols were presented, the subjects reported correctly; however, from 5 symbols on, subjects overestimated. At about 18 symbols, estimates started getting lower until at 22 symbols (for 500 ms) the median estimate was close to the true amount. At 22 symbols at 200 ms exposure time, the median estimate was even lower and, in fact, represented an underestimate of the true symbol amount. Previous studies using dots as stimulus material (Kaufman et al, 1949; Mintern and Reese, 1951; Taves, 1941) showed similar functions when median estimated number was plotted against presented number, except that estimates were more accurate and generally lower than in the present study. In the Kaufman study, using dots as the stimuli, up to 5 were reported correctly, from 6 to 9 dots the reports were overestimates, and above 10 dots the reports were underestimates. Although visual angle

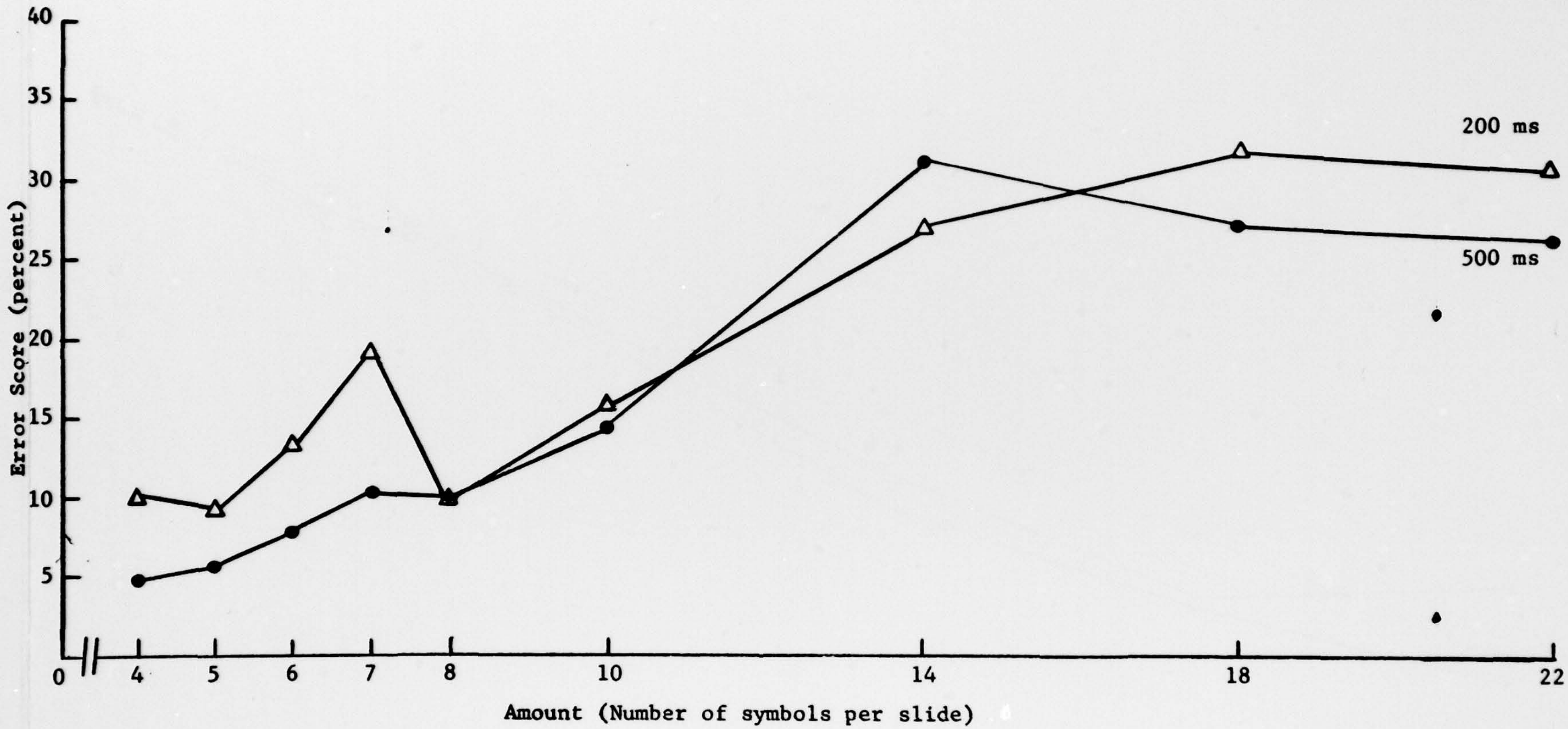


Figure 1. Mean Error Score at Each Amount Level for Two Exposure Times.

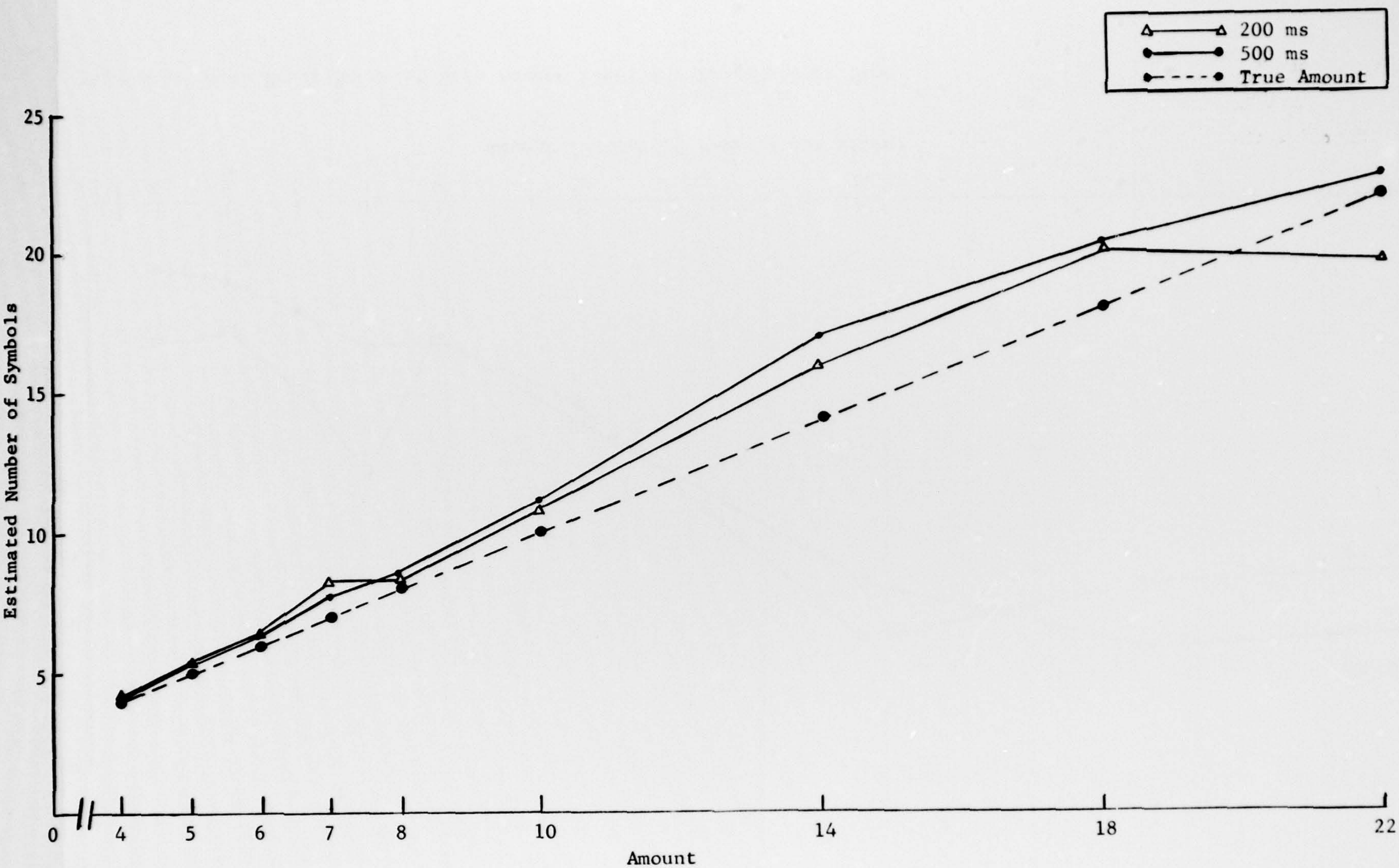


Figure 2. Median Number of Symbols Estimated at Each Amount Level for Two Exposure Times.

was not stated in the Kaufman (1949) study, it was possible to estimate from information given the approximate range of visual angles subtended by stimulus materials. These ranged from 2° with the lower numbers of dots to approximately 10° with the larger numbers. The visual angle subtended by the visual image in the present study was considerably larger (22°); therefore, it was not unreasonable to expect the present task to be more difficult. However, the similarity of the findings in spite of the stimulus differences is striking. This similarity is also reflected in percent error data; in previous studies (Mintern and Reese, 1951; Taves, 1941; and Kaufman et al, 1949) and in the present study, error increased in much the same manner as a function of the increase in amount presented.

Along with this error increase, Taves (1941), Mintern and Reese (1951), and Kaufman et al (1949) found an increase in variability of response with an increase in amount presented. Although military symbols were used in the present study and dots in the previous studies, a similarity exists in the relationship between variability in estimated number of symbols and number presented in the present study and in earlier studies mentioned above. Figure 3 shows that variability increased slowly up to approximately 8 symbols and then more rapidly to 22 symbols. This is typical of previous findings.

The subjects in the present study later took part in an experiment in which they were to recall an aspect of the perceptual organization of the military symbols used in the present study. The symbols were presented for one minute. A more detailed account of the experiment can be found in Ringel and Vicino (1964). This circumstance permitted comparison of performance by the same subjects in three perceptual tasks, two tachistoscopic and one nontachistoscopic. Spearman's rank order correlation based on error scores was calculated between rankings at the two exposure times in the tachistoscopic study and between each of these rankings and the ranking of the same subjects on the nontachistoscopic task. For performance at 200 ms, correlation between the tachistoscopic and the nontachistoscopic tasks was $r = .31$; at 500 ms, it was $r = .43$. Both correlation coefficients were rather low considering the similarity in stimulus material and ambient conditions. The probability of predicting relative success from one task to the other is low enough to introduce considerable caution in generalizing from the results of a tachistoscopic study to one where more time is allowed to view the stimulus materials. The rank order correlation coefficient between tachistoscopic tasks at 500 and 200 ms was .58--again rather low for tasks with the same stimulus material, the same ambient conditions and, in this case, the same task.

EXPOSURE TIME AND PERFORMANCE

Significant differences were found for the effects of exposure time on error score. Higher percent error (Figure 1) and greater variability in estimated number of symbols (Figure 3) associated with the 200 ms condition than with the 500 ms condition were found and, as mentioned above, correlation between performance at the two time conditions was not as high as expected.

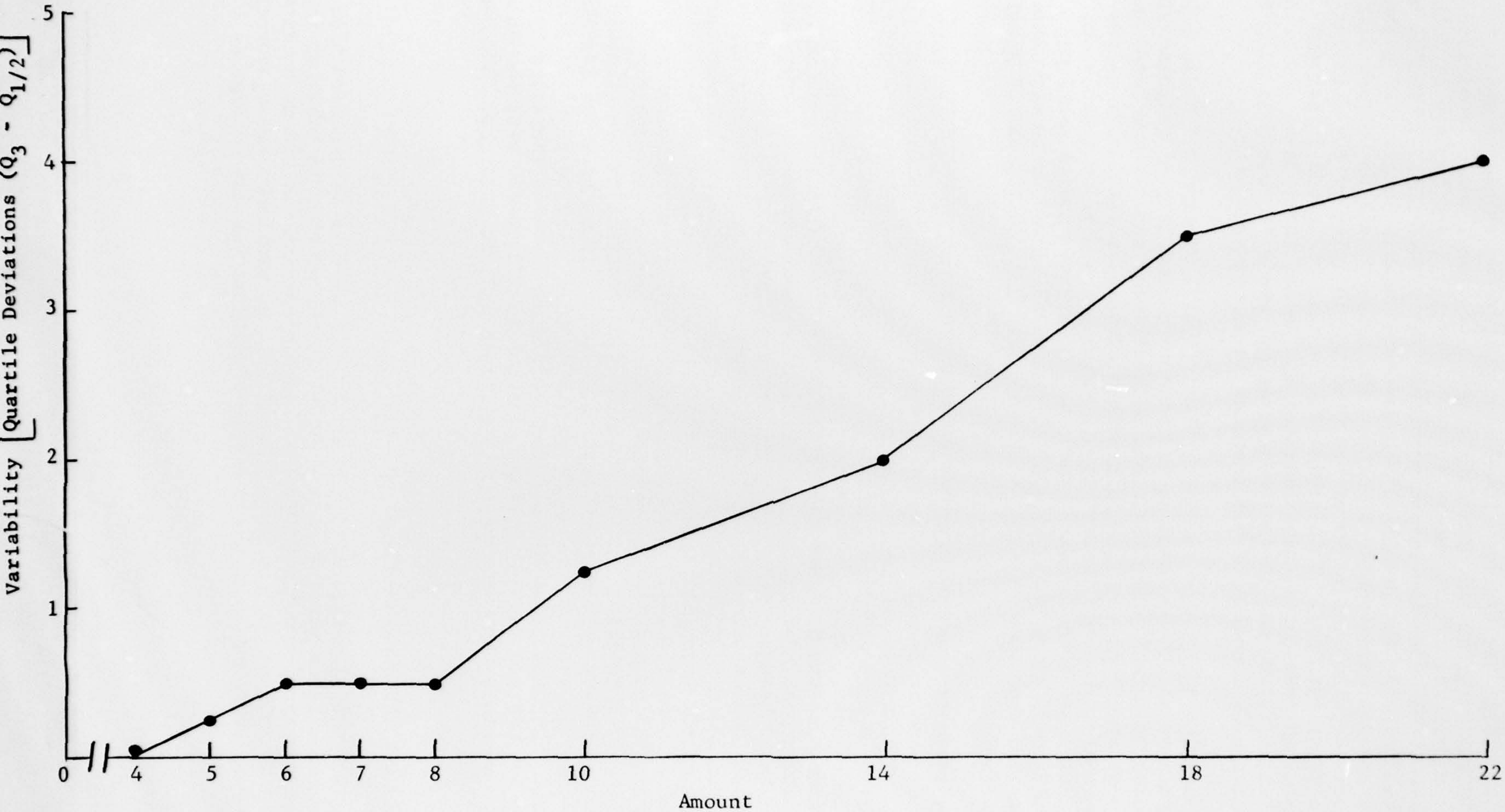


Figure 3. Variability in Estimated Number of Symbols as a Function of Amount.

One possible explanation for the differences in error score at the two exposure times is the same as that advanced earlier to explain irregularity in the error score curve. Approximately 200 ms is the average reaction time of the eyes in shifting from one fixation point to another. Therefore, with time for only one fixation, there can be only one "act of attention" and serial counting cannot occur. The 500 ms exposure allows for two fixations, more of the slide area may be exposed to foveal vision, and some form of counting may begin. As shown in Figure 1, the percent error score was considerably greater at 200 ms than at 500 ms for up to approximately 7 symbols. Beyond 7 symbols, the task was difficult enough so that two fixations did not enhance the performance at 500 ms with any consistency. Although at the 18- and 22-symbol level, a more realistic estimate of the number of symbols was made at the longer exposure time, the difference was relatively small in terms of total error. In the present task, the visual angle was approximately 22°. The present analysis led to the hypothesis that the differences between 200 ms and 500 ms would not occur if the visual angle of the display were 2° or less, which would allow for foveal vision so that no more than one fixation would be needed. The test of this hypothesis awaits an experiment where visual angle and exposure time are varied and discrimination accuracy measured for all possible combinations.

EXPOSURE TIME ORDER AND PERFORMANCE

There were no significant differences as a function of order of exposure time (Table 1).

Allowing the subjects to respond to the stimuli at the 500 ms exposure first did not affect error score any differently than allowing them to respond first at 200 ms. Practice effects were therefore virtually nonexistent for the discrimination of visual number in the present tachistoscopic task.

CERTITUDE AND NUMEROUSNESS

A three-way analysis of variance with amount, exposure time, and order of exposure time as the treatment classifications and certitude as the dependent variable was performed on the data. Ordinal values of 1 to 5, "absolutely uncertain" to "absolutely certain", were used in the analysis since the data could not be fitted to the successive category scaling model, thus precluding a determination of scale values for the intervals. The summary of the analysis of variance shown in Table 2 reveals order of exposure time as the only main effect which was not statistically significant. In addition to the significant amount and exposure time effects, there was one significant interaction, amount x time. The nature and magnitude of these effects are more easily discernible in Figure 4.

Table 2

ANALYSIS OF VARIANCE FOR CERTITUDE JUDGMENTS

Source of Variation	d.f.	Mean Square	F	P
Order of presentation (O)	1	6.46	.96	N.S.
Btw Subjects in order S (O)	30	6.76		
Exposure Time (T)	1	10.84	12.81	.01
T x O	1	.50	.59	N.S.
T x S (O)	30			
Amount (Number of symbols) (A)	8	64.45	97.16	.01
A x O	8	.23	.34	N.S.
A x S (O)	240	.66		
A x T	8	1.05	4.04	.01
A x T x O	8	.42	1.61	N.S.
A x T x S (O)	240	.26		
TOTAL	575			

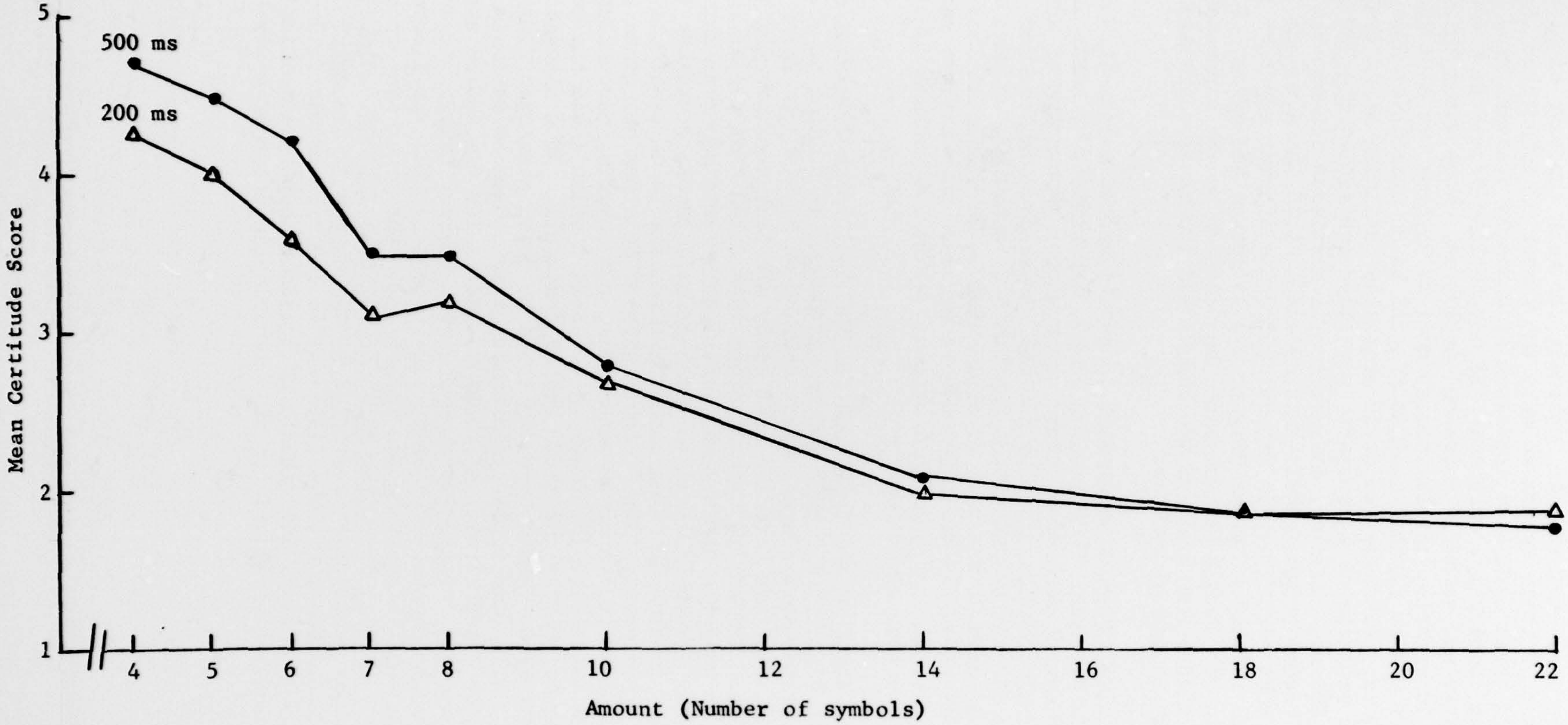


Figure 4. Mean Certitude as a Function of Number of Symbols Displayed for Two Exposure Times.

The shape of the function for mean certitude at 200 ms and 500 ms exposure times are very similar. As expected, for each exposure time, subjects were most certain of their responses when the number of symbols on the slide was smallest and their certitude decreased at a nearly constant rate up to 7 symbols. Then, curiously, certitude did not decline from 7 to 8 symbols at 500 ms exposure and actually increased at 200 ms. More about this later. From 7 through 18 symbols, certitude continued to decline but at a decreasing rate. There was no further decline from 18 to 22 symbols at 200 ms; however, at 500 ms, certitude was still gradually declining at 22 symbols. For all amounts starting at 4, subjects were more certain at the 500 ms exposure than at 200 ms. However, differences in certitude between the two exposure times became progressively smaller until at 18 there was no difference and at 22 more certainty was expressed for the shorter exposure time. Plotting medians instead of means (Figure 5) to reduce the effect of extreme scores did not change the shape of the function to any appreciable degree, except that coincidence of the 200 ms and 500 ms exposure times occurred at both 14 and 18 symbols.

Amount by time interaction appears to have occurred mainly at the 18- and 22-symbol conditions. This finding is not explainable by performance, since at the 22-symbol condition, there was a higher percentage of error for the 200 ms exposure than for the 500 ms. A possible explanation, based on the tendency for the two certitude functions to converge, is that beyond a given number of symbols the task of estimation is so difficult that the small benefits attributable to difference in exposure time are not reflected in confidence. Beyond that point one exposure would not produce consistently greater confidence than the other.

The general shape of the functions for the 8- through 22-symbol conditions shown in Figure 5 for both 200 ms and 500 ms is similar to results reported in the literature (Taves, 1941; Kaufman et al, 1949). Below 8 symbols, however, findings from the APRO study are less consistent with previous findings. Typically, experimenters have found that median certitude hovers at or near the upper limit (absolutely certain) up to about 5 stimuli. From the 5-stimuli level to the 6-stimuli level, certitude then drops off sharply in much the same manner as from the 3- to 4-stimuli level in the present experiment. This discontinuity of function has given rise to the postulation of two separate mechanisms, one covering perception of numerousness and associated certitude up to about 5 stimuli and the other numerousness and associated certitude beyond 5 stimuli. If a discontinuity of function parallel to that previously found is to be extracted from the present data, it will have to be via extrapolation. Such an estimate would place the point of discontinuity at about 3 symbols for the 500 ms exposure and at about 2 symbols for the 200 ms exposures, thereby tending to confirm the notion of two perceptual mechanisms but indicating that the discontinuity demarcating the change from one mechanism to the other can occur at other than 5 or 6 stimuli depending on the conditions imposed. Since the certitude functions pretty much conformed to the shapes of the error functions, the probable reason for the disparity in location of the discontinuity between APRO findings and results of previous certitude studies is found in the explanation adduced for the disparity in error of estimate functions.

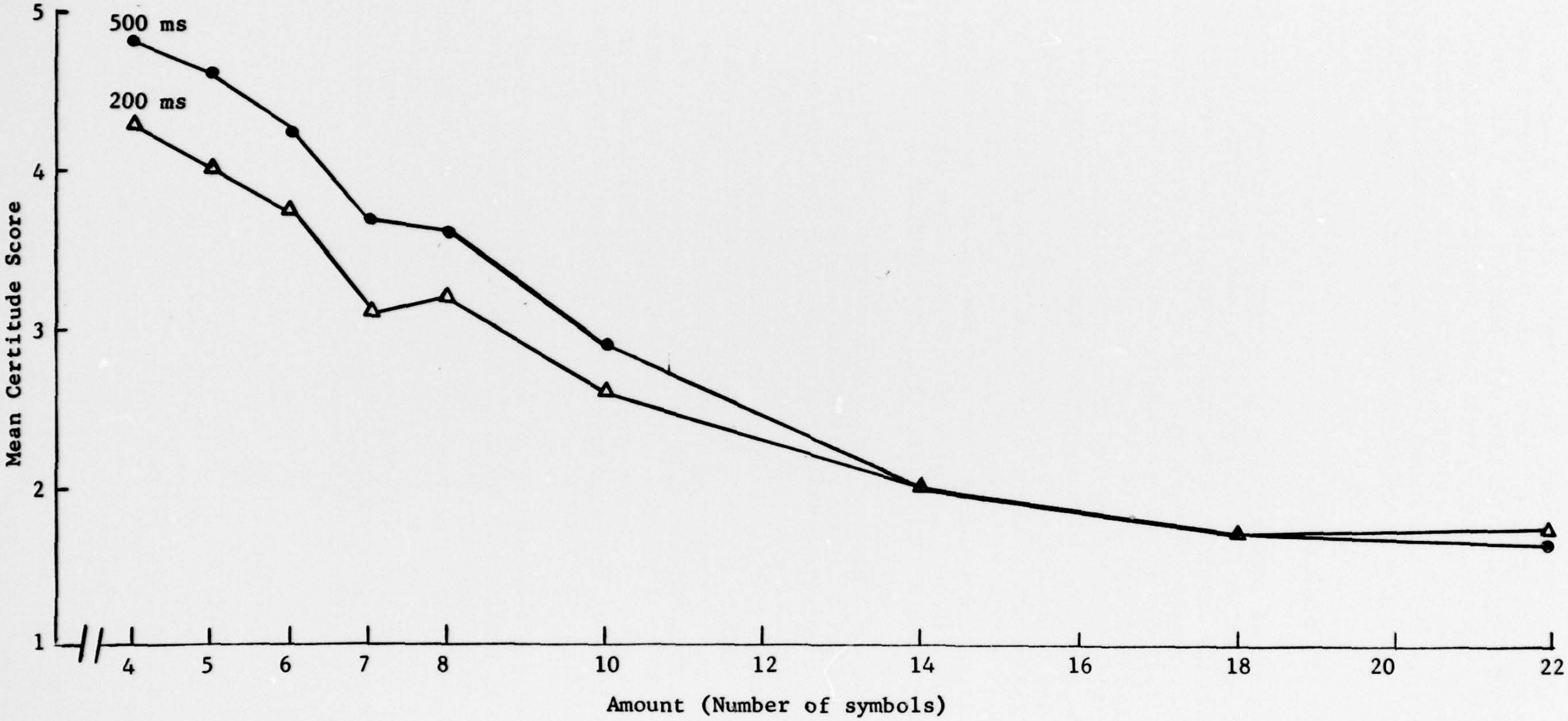


Figure 5. Median Certitude as a Function of Number of Symbols Displayed for Two Exposure Times.

While Figures 4 and 5 seem to offer evidence of another discontinuity at about the 7-symbol level, the irregularity is unlike any discontinuity found in previous studies. That it is really a discontinuity is doubtful since if the 7-symbol condition were omitted, the total function from 4 through 22 symbols would be relatively smooth and continuous. Again, the probable reason for this seeming discontinuity is the similar apparent discontinuity in the error score, reason for which has been exposited. The same reason probably accounts for the inflection between 18 and 22 symbols.

CERTITUDE AND ERROR SCORE

The general conformity of the certitude function to the error score function may belie the relationship between the individual's perception of number and his certitude about the accuracy of his perception in terms of the product moment correlation. Based on 288 observations (9 for each of 32 subjects), the correlation coefficient obtained between certitude and error score was -.30 at the 500 ms exposure, and -.35 for the same number of observations at 200 ms exposure, indicating that as error increased, certitude decreased. Yet when the between people effects (tendency for correlation to be inflated or depressed because of individual differences) were partialed out of these correlation coefficients, that portion of the correlation attributable to within people effects produced an r of -.88 for the 500 ms exposure and an r of -.58 for the 200 ms exposure. Apparently, shorter time exposure had more of an effect on a man's perception of his performance than it had on his actual performance.

Within the present study, a good range of difficulty was represented by the conditions imposed: numerousness levels covered the most critical range, judging from previous studies and from the numbers of symbols that might be displayed simultaneously in an operations center. In addition, the performance-certitude frequency distributions at 200 ms exposure time were quite similar to those at 500 ms. It would thus appear reasonable to conclude that the degree of relationship between performance and confidence in that performance can vary considerably depending upon the conditions imposed (500 ms vs 200 ms). There also appears to be considerable variation across subjects in the certitude they assign to a given performance level, judging by the magnitude of the increase in correlation when between people effects were removed.

No direct treatment of the certitude-performance relationship was found in reported research in which discrimination of visual number and certitude measures were taken. However, the study which followed the present one did involve such comparisons (Andrews and Ringel, 1964). The task was different and stimuli were not presented tachistoscopically, but since both studies involved extraction of information from displays and subjective feelings of certitude about the accuracy of the extraction, the similarities and dissimilarities in the findings are of interest.

Correlation of $r = .-.55$ between error and certitude was obtained for the within person effects in the Andrews and Ringel study compared to $r = -.88$ and $r = -.58$ obtained in the present study. The between person effects obtained were $r = .50$ (Andrews and Ringel) and $r = .18$ and $r = -.22$ (present study). The contrasting results seem to indicate that people differences can vary widely depending on the nature of the task and difficulty levels within a task.

The relationship can be looked at another way. At the 500 ms exposure, there were 60 "absolutely certain" responses, of which 78% were, in fact, correct. There were 109 correct responses, only 43% of which received the "absolutely certain" response. At the 200 ms exposure, there were 32 "absolutely certain" responses, of which 59% were actually correct, and 86 correct responses, only 22% of which received the "absolutely certain" response. These figures indicate greater probability that a man will be right in his estimate when he says he is "absolutely certain" than there is that he will say he is "absolutely certain" when, in fact, he is right. This is logical since on many occasions a man knows his answer is only an approximation, albeit a close approximation, and thus is not absolutely certain though in fact he may be right. So, too, with outright guesses. The relationship described above holds for both the 200 ms and 500 ms exposures, though a man is more likely to be wrong when he says "absolutely certain" and less likely to say "absolutely certain" when he is right at the 200 ms exposure than at the 500 ms.

Addressing the question whether individuals tend to have a pattern of certitude consistently higher or lower than other individuals irrespective of their actual performance on the task, Spearman rank order correlations based on mean certitude rank for the 200 ms exposure, the 500 ms exposure, and the nontachistoscopic presentation were computed. The results are as follows:

$$r_{200\text{ ms } 500\text{ ms}} = .77$$

$$r_{200\text{ ms NT.}} = .54$$

$$r_{500\text{ ms NT.}} = .31$$

Only the last of these is not significant at the .05 level. As would be expected, highest correlation was between the two tachistoscopic tasks. Not so expected was the higher relationship of the 200 ms exposure than the 500 ms to the nontachistoscopic task, particularly since the reverse was true in terms of actual error. It would seem that to the extent these differences are not attributable to chance, there are more aspects in common to the 200 ms and nontachistoscopic tasks, apart from accuracy of performance, which contribute to feelings of certitude. While there may be some tendency for one individual to be consistently more or less certain than others, the magnitude of this relationship apparently can vary considerably across differing tasks.

THE CERTITUDE CONTINUUM

One final aspect that warrants at least a cursory look is the response continuum itself. As mentioned previously, data from the five-interval certitude continuum were not scalable by the method of successive categories. Why this was so for this continuum and not for the certitude continuum used in the Andrews and Ringel study (1964) is difficult to ascertain. Apparently, there was more nonrandom error in the present study. Whether this is attributable to differences in the nature of the task, differences in the continuum (eight-category fully anchored vs five-category end anchored), some combination of these two factors, or other unspecified causes, could not be determined. Nor does the literature report any scaling data on the certitude continuums used. In information theory parlance, 2.28 "bits" of response information were provided by the five-category continuum and 2.85 "bits" by the eight-category continuum. In amount of information provided as a proportion of the amount that could be provided by continuums of these lengths (relative entropy), $R_e = .98$ for the five-category scale and $R_e = .95$ for the eight-category continuum. Bendig and Hughes (1953), in a study of the effects of anchoring and number of scale categories on transmitted information, found that their five-category scale produced 2.22 bits of information and an $R_e = .96$. Thus, the five-category continuum used in the present study appears not to be deficient in terms of "bits" of response information. Certainly, no fewer than five categories are indicated for judgments of this type. More than five categories could probably be used effectively, since the five-category scale produced near the upper limit of response information possible for the length of scale, and the eight-category scale in the other study was also producing close to the maximum possible. While there is nothing in the evidence to recommend the five-category continuum over the eight, there is at least probable evidence to recommend the eight- over the five-category continuum.

IMPLICATIONS OF FINDINGS

By intent, the present study is of value chiefly in terms of the implicative rather than definitive nature of the results obtained. First, discrimination of visual numerousness functions appears to be relatively immune to the form characteristics of the stimulus--dots, circles, military symbols, etc. The implication is that the perceptual mechanisms reported in the literature for less meaningful stimuli can be postulated for symbolic displays of military content. Similarly, the shape of these functions is relatively unaffected by the visual angle (involving other than foveal vision) subtended by the display area or by differences in exposure time as represented by two common tachistoscopic exposure times. Again, the implication is that the previously reported mechanisms operate with visual angles approaching those expected in command systems displays and without regard for whether exposure time permits a saccadic eye movement, as long as the time is sufficient to subitize or to estimate but not to count.

When the interest of an experiment is in some absolute measure of performance--accuracy, error rate, etc.--rather than general shape of the performance function, it is important to treat each new or changed variable or condition imposed as a different entity whose specific values are yet to be determined. This is implied from the finding that error or accuracy can vary as a function of time and perhaps visual angle even though the nature of these relationships to numerousness is highly similar.

The foregoing implications also apply to certitude as used in this study. An additional implication which is derived from direct correlational analysis of the certitude-accuracy relationship is that degree of relationship can vary widely as a function of the task imposed and is not necessarily associated with the apparent similarity of the tasks. Thus, the certitude-accuracy relationship is not generalizable but must be determined separately for each task which differs in any dimension from any previously analyzed task. It cannot be assumed that certitude response continuums alleged to be adequate in previous studies without quantitative analysis or determined merely on some logical basis will have the desired psychometric, information transmittal, or sensitivity properties. Such assumptions could severely limit the information obtained and the statistical manipulations permitted.

There are a number of implications for further research but these are more germane to basic questions left unanswered in the experimental literature on visual processes than they are to the operationally oriented program of research planned in the Command Systems Task. For example, a definitive study on the effects of visual angle, to include both foveal and peripheral vision, on discrimination of visual number would answer questions raised but left unanswered by the present study. Closely related would be an investigation of the effects of average deviation of stimulus figures from the focal point in a slide (or from foveal vision) at various levels of numerousness. Saccadic eye movement and discrimination as a function of varying time exposure would also bear scrutiny.

REFERENCES

Andrews, R. S. and Ringel, S. Certitude judgments and accuracy of information assimilation in visual displays. Technical Research Note 145. May 1964.

Bendig, A. W. and Hughes, J. B. Effects of amount of verbal anchoring and number of rating scale categories upon transmitted information. Journal of Experimental Psychology, 146, 1953, 87-90.

Kaufman, E. L., Lord, M. W., Reese, T. W., and Volkman, J. The discrimination of visual number. American Journal of Psychology, 1949, pp 498-525.

Minturn, A. L. and Reese, T. W. The effect of differential reinforcement on the discrimination of visual number. American Journal of Psychology, 1951, 31, pp 201-231.

Ringel, S. and Vicino, F. L. Information assimilation from symbolic displays--Amount of information presented and removed. Technical Research Note 139. March 1964.

Saltzman, I. J. and Garner, W. R. Reaction time as a measure of span of attention. American Journal of Psychology, 1948, 25, pp 227-241.

Taves, E. H. Two mechanisms for the perception of visual numerousness. Archives of Psychology, 1941, pp 5-46.

Woodworth, R. S. and Schlosberg, H. Experimental Psychology. New York: Henry Holt and Company, 1956.